A Fusion Framework with Nonlinear Degradation Improvement for Remaining Useful Life Estimation of Lithium-ion Batteries

Datong Liu, Limeng Guo, Jingyue Pang, Yu Peng

Department of Automatic Test and Control, Harbin Institute of Technology, Harbin, 150080, China

liudatong@hit.edu.cn
glm_hit1991@163.com
pjy19880909@163.com
pengyu@hit.edu.cn

ABSTRACT

Fusion prognostic framework for lithium-ion battery remaining useful life (RUL) estimation has become a hot spot. Especially, the cycle life prediction has been conducted widely, for which many prognostic methods have been proposed correspondingly. However, many fusion frameworks which can achieve high precision are accompanied with high computing complexity and high time consumption which makes these methods low real-time performance. Either, some widely used prediction models with low complexity are weak to handle the nonlinear degradation features. To solve these problems, a fusion framework is proposed combining the model-based extended kalman filter (EKF) and the data-driven improved nonlinear scale degradation parameter based autoregressive (NSDP-AR) models. The proposed approach takes advantage of the state tracking ability of EKF algorithm to define the specific state transition model for the battery sample. Meanwhile, NSDP-AR model which contains the degradation features of each period is to promote the universality of the ND-AR (Nonlinear Degradation Autoregressive) model. NSDP-AR model is used to obtain the long term trend prediction results which are adopted as the observation data. Finally, a combination is made to realize the RUL prediction under the kalman filter (KF) system, which is an improvement to meet the practical applications. Experimental results with the battery test data from NASA PCoE and CALCE show that the fusion prognostic framework can predict the lithium-ion battery RUL with high efficiency and accuracy.

1. INTRODUCTION

Lithium-ion (Li-ion) batteries have become the preferred energy solution for various electrical-driven systems such as consuming electronics, electric vehicles, and even the aerospace field due to the high energy density, high galvanic potential, lightness of the weight and long lifetime compared to traditional energy storage batteries (He, Williard, Osterman & Michael Pecht, 2011). However, no matter how excellent the performance of the lithium-ion battery is, it degrades over time for aging, environmental impacts, and dynamic loading (Zhang & Lee, 2011). In order to satisfy the increasing demand for operation reliability, study on the effective methodologies for battery performance evaluation becomes considerably necessary and important. In particular, the remaining useful life (RUL) estimation of the Li-ion battery is the essential part in the field of electronic prognostics and health management (PHM). RUL also can be named remaining service life or residue life, which refers to the available service time (always using how many charge and discharge cycles the battery can experience to describe this variable) left before the degradation level of the system is unacceptable (Zhang & Lee, 2011). Successful RUL prediction is highly desirable for ensuring reliable system operation.

Recently, extensive research activities have been conducted on the RUL estimation of Li-ion batteries. Generally speaking, prognostics methods can be classified into data-driven and model-based approaches. Artificial Neural Networks (ANN) (Liu, Saxena, Goebel, Saha & Wang, 2010) and Relevance Vector Machine (RVM) (Zhang & Lee, 2011) are typical representatives for data-driven approaches which establish the prediction model using the characters selected from the data without considering the physical system features. Model-based approaches focus on the state space model studies which proceed from the system characteristics like Extended Kalman Filter (EKF) (He, Williard, Osterman & Pecht, 2011) and Particle Filter (PF) which can obtain both the RUL and uncertainty representation (Saha & Goebel, 2009).

Current researches show that there are several limitations in the above two kinds. Although the empirical degradation model is proposed on the basis of a large amount of experiments and rigorous analyses, it still cannot take into account all the complicated online operation conditions. That is to say, the performance of model-based method which is determined by the accuracy of the model cannot adapt to various actual working conditions. Meanwhile, estimation of the parameter in a specific model is an uneasy work and satisfied results are rarely obtained. Even though there is little such problem in data-driven approaches, they still suffer from the drawback for ignoring the distinctions among different types of systems in which the data sets belong to.

In order to address the aforementioned problems, data-model-fusion prognostic frameworks are proposed and attract more and more attention. Till now, various studies have been conducted in this field. A fusion prognostic framework with model-based PF algorithm and data-driven neural networks method has been proposed by Liu etc for dynamic system state forecasting and achieved a successful result for battery RUL estimation (Liu, Wang, Ma, Yang & Yang, 2012). Fusion frameworks combining data-driven and model-based prediction approaches can overcome the aforementioned shortcomings. With adoption of the data features obtained from the data-driven approach, dependence of the model-based method on the empirical model is alleviated. Meanwhile, system characteristics are added into the data-driven framework. In short, both the system and data features are contained in the fusion framework whose effect is expected to be better than any single method.

For online real-time applications, the accuracy as well as the real-time performance of the prognostic algorithm is crucial. However, majority of the fusion framework focuses on the accuracy without considering the efficiency and the calculation complexity. In order to satisfy the rigorous requirements in practical forecasting, a data-model-fusion framework with a strong emphasis on real-time prediction capability is worthy of more studies.

EKF algorithm used for the nonlinear systems is an extension of the Kalman Filter (KF) which is a recursive solution under the least-squares principle. EKF is one kind of stochastic filtering method based on the state space model which stands for the system features. With linearization of the system equations using Taylor expansion and cycles of state estimation and updating, EKF can provide an efficient computational solution even when the precise nature of the modeled system is unknown (Welch & Bishop, 1995). EKF has a good state tracking ability in the capacity prediction and the RUL can be obtained when the failure threshold is provided (He, Williard, Osterman & Pecht, 2011). Meanwhile, massive applications of the performance estimation for the lithium-ion batteries show that EKF is a promising algorithm with low computational complexity, strong real-time estimation ability and satisfied prognostic effect (Zhang & Lee, 2011).

Although such advantages have been shown in practical applications, EKF still has the limitation as a model-based approach. As discussed before, adoption of a data-driven method can effectively address this problem. Various data-driven methods such as machine learning and artificial intelligence have a satisfied prediction effect but also with a complicated calculation process and a lot of time consumption which make these algorithms have less practical value in real-time battery performance estimation especially when implemented in hardware. Due to such a practical concern, a much simpler data-driven method that has been applied in many prediction fields, namely the AR (Auto Regressive) model, draws our attention. AR model is suitable for the real-time estimation with small data sets (Wei, 1994). This combination is expected to possess better efficiency for real-time prognostic applications.

However, degradation of the Li-ion battery has obvious nonlinear characteristics. As a consequence, linearization using Taylor expansion leads to inaccurate approximation of state transition and observation equations resulting in non-optimal battery performance estimation. Meanwhile, AR model is a linear model which establishes a linear equation between the current state and several previous states. Although researchers have conducted some studies on the AR model improvement (Liu, Luo, Peng, Peng & Pecht, 2012), there still exists certain theoretical and application limitations in the modified AR model.

To address these problems, a data-model-fusion prognostic framework is proposed by combining the EKF algorithm with a modified nonlinear scale degradation parameter based autoregressive (NSDP-AR) model for the RUL estimation. EKF system is established on the basis of the state space model which contains state transition and observation functions. Using the strong state tracking ability of the EKF algorithm, parameters in the empirical degradation model (i.e. the state transition function) (Saha & Goebel, 2009) are obtained so that a specific model for the battery sample is established. To improve the absolute degradation parameter based ND-AR model for better long term prediction, a more reasonable scale parameter based nonlinear degradation factor is proposed which contains the degradation changing information for certain type of batteries. Correspondingly, NSDP-AR model is established by combining the proposed factor and the AR model. For a reasonable extension of the obtained NSDP-AR model in practical applications, Grey Correlation Analysis (GCA) method is used to determine the weight of the parameter groups from different battery samples which are used to establish the modified model using the true degradation information. On the basis of those parameters and corresponding weights, specific NSDP-AR model is obtained for the battery's RUL estimation. On the basis of the above work, fusion-data-model prognostic framework is established. Experimental results using the NASA PCoE and CALCE battery data sets show that the framework can predict the Li-ion battery RUL efficiently and accurately which indicates a strong practical application of the proposed framework. While, there is one thing that need to be pointed out that there are many reasons that will cause the degradation of the battery. Here, we only consider the major part of these factors, which are the battery aging caused by charge and discharge cycle. We only take the cycle aging caused degradation into consideration here.

This paper is organized as follows. In Section 2, the related prediction models including KF/EKF algorithm and current ND-AR model as well as the correlation analysis method GCA are introduced. The proposed NSDP-AR model and the corresponding fusion-data-model prognostic framework for RUL estimation of the Li-ion battery are introduced in Section 3. The effectiveness of the proposed prognostics framework is demonstrated via battery RUL prediction experiments using Li-ion battery data sets from NASA PCoE and CALCE in Section 4. Finally, the conclusion and future work are given in Section 5 and 6, respectively.

2. Related Work

2.1. KF/EKF Algorithm

Many researchers have used KF/EKF algorithm to estimate the unknown parameters in the battery empirical degradation model and obtain the RUL of the battery (He, Williard, Osterman & Pecht, 2011). These researches used the state tracking ability of KF/EKF algorithm and the state space model of the battery system. EKF algorithm is the expansion of the KF algorithm in order to meet the requirements of the nonlinear applications. So, here we give more information of KF algorithm, EKF theory can be obtained by analogy of the KF algorithm.

KF is one kind of model-based algorithms that provides an efficient recursive solution of the least-squares method for the discrete-data linear systems. The state space model of the system can be described in Eq. (1).

$$\begin{cases} x_k = F_k x_{k-1} + B_k u_k + w_k \\ z_k = H_k x_k + v_k \end{cases} \tag{1}$$

Here, F_k stands for the state transition matrix, B_k is the control matrix, x_k is the k^{th} state of the system, u_k is the control input of the system, while the w_k is the system process noise which obeys the Gaussian distribution that its mean value is zero and the variance is W. The first equation named state transition function describes the relationship between the k^{th} (i.e. x_k) and the k-1th (i.e. x_{k-1}) state. The second equation named observation function turns the implied system state into measurable outputs. Here, the H_k is the observation matrix, z_k is the measurements and the v_k is similar with w_k but its variance is R where W and R are both real.

The prediction process using the KF algorithm is divided into two main steps: time update and measurement update which can be described by two groups of mathematical equations (Welch & Bishop, 1995). The calculation flow of KF algorithm is as Figure 1 shows.

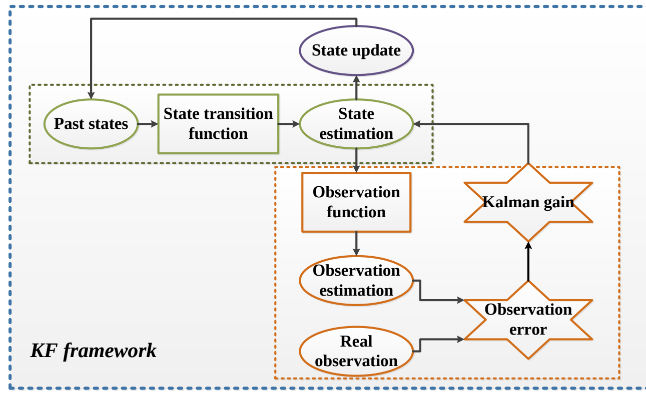

Figure 1. KF algorithm framework

Time Update:

$$\begin{cases} x_{k|k-1} = F_k x_{k-1|k-1} + B_k u_k \\ P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \end{cases} \tag{2}$$

Measurement Update:

$$\begin{cases} y_k = z_k - H_k x_{k|k-1} \\ S_k = H_k P_{k|k-1} H_k^T + R_k \\ K_k = P_{k|k-1} H_k^T S_k^{-1} \\ x_{k|k} = x_{k|k-1} + K_k y_k \\ P_{k|k} = (I - K_k H_k) P_{k|k-1} \end{cases} \tag{3}$$

Here, the meaning of F_k , B_k and u_k are same with the introduction in Eq. (1). $x_{k|k-1}$ and $P_{k|k-1}$ are the predicted mean and covariance of the state, respectively, on the time step k before seeing the measurement. $x_{k|k}$ and $P_{k|k}$ are the estimated mean and covariance of the state, respectively, on time step k after seeing the measurement. y_k is the innovation or the measurement residual on time step k. S_k is the measurement prediction covariance on the time step k. K_k is the filter gain, which tells how much the predictions should be corrected on time step k. Time update step depends on the k-1th state measurement update result $x_{k-1|k-1}$ and $P_{k-1|k-1}$ as well as the k^{th} control input of the system. At the same time, on the basis of the measurement residual, the filter gain can be obtained to update the previous estimation result in Eq. (2). This step means using the real measurement information to update the previous result. After each time and measurement update as shown in Eq. (2) and Eq. (3), the process is repeated to obtain the state estimate and corresponding prediction covariance. More detailed description of these equations and theory can be referred in related literature (Welch & Bishop, 1995).

However, KF algorithm is used for the systems whose features can be described into a linear equation as shown in Eq. (1). In most of the practical applications like Li-ion battery, systems always have some nonlinear characteristics such that the system model has to be described as follows, the meaning of the variables are same with Eq. (1):

$$\begin{cases} x_k = f(x_{k-1}, u_k, w_k) \\ z_k = h(x_k, v_k) \end{cases} \tag{4}$$

In order to extend KF algorithm into nonlinear conditions, EKF which uses the Taylor expansion to linearize the system functions around the current estimate is employed. After the linearization step, the same approaches as KF algorithm are used to estimate the system states. EKF is a subprime optimal estimate method which can satisfy the requirements of certain systems which have a low or median nonlinear feature.

2.2. The ND-AR Model

To address the nonlinear degradation prediction problem, an improved AR model named ND-AR model which describes the degradation features using an accelerated factor K_T as follows:

$$K_T = \frac{1}{1 + a \cdot (k + b)} \tag{5}$$

Here, a and b stand for the unknown parameters in the factor which contain the nonlinear degradation information of the capacity, and k is the prediction step which can be described the discharging cycle number the battery has experienced in another way. Here, the ND-AR model only considers the battery aging caused by charge and discharge cycle which is the major reason of the degradation. This means that ND-AR model chooses an approximated way to describe the nonlinear feature of the capacity through only taking the cycle aging into consideration but ignoring other factors. The factor extracts a correlation relationship between the prediction step k and the accelerated feature of the capacity degradation. This improvement provides a valuable reference that adopts a mathematical factor related to the degradation parameters to describe the degradation characteristics in the data-driven prediction model. The specific expression of the ND-AR model is as follows:

$$x_t = K_T \times [\phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + a_t] \tag{6}$$

Here, K_T is the accelerated factor and the remaining part of the Eq. (6) is the basic AR model. In the AR model, ϕ_k is the autoregressive coefficient and a_t is the noise which obeys Gaussian distribution. The adoption of the accelerated factor makes the AR model nonlinear. As a result, the accelerated and nonlinear degradation trend of the capacity of Li-ion battery can be retrieved (Liu, Luo, Peng & Pecht, 2012).

2.3. GCA Algorithm

GCA is an efficient and convenient correlation analysis approach which has been generally used in data variation trend similarity analysis. After choosing the standard data array $Y=\{Y(k) \mid k = 1,2, \dots,n\}$ and the comparison data set $X_i=\{X_i(k) \mid k = 1,2,\dots,n\}$ $(i = 1,2,\dots,m)$, correlation coefficient can be calculated according to the following equation:

$$\varsigma_i(k) = \frac{\min_i \min_k |Y(k) - X_i(k)| + \rho \max_i \max_k |Y(k) - X_i(k)|}{|Y(k) - X_i(k)| + \rho \max_i \max_k |Y(k) - X_i(k)|} \tag{7}$$

where ρ is the discrimination whose value ranges from zero to infinity. According to some investigation, the effect of the resolution is best when $\rho \le 0.5463$ (Shen, Xue & Zhang, 2003). In order to facilitate the overall comparison, a concentration which turns the correlation coefficient of each point into one final result will be made through calculating the average of the above coefficients. The correlation degree between Y and X_i is defined in Eq. (8):

$$r_i = \frac{1}{n} \sum_{k=1}^{n} \varsigma_i(k), k = 1, 2, ..., n \tag{8}$$

Here, ς_i is the correlation degree between X_i and Y and n is the number of the comparison data sets.

3. Fusion Prognostic Framework for Lithium-ion Battery RUL Estimation

3.1. Estimation of the Empirical Degradation Model

In order to establish the specific state space model for the Li-ion battery, we need to confirm the state transition function which indicates the system state transition features. Researchers such as Saha in NASA PCoE have conducted studies on establishing an empirical state model for lithium-ion batteries and proposed a degradation model to describe the capacity degradation properties (Saha & Goebel, 2009). The model can be described as follows:

$$C_{k+1} = \eta_C C_k + \beta_1 e^{(-\beta_2/\Delta t_k)} \tag{9}$$

The model makes a relevancy between the k^{th} charge cycle capacity C_k and the $k+1^{th}$ discharge cycle capacity C_{k+1}. Moreover, Δt_k is the rest time between cycles k and $k+1$, η_C is the charge and discharge efficiency named coulomb efficiency which describes the difference between the capacity filled in the battery and the capacity that the battery could provide during usage, β_1 and β_2 are the parameters that need to be determined. The first part of the model $\eta_C C_k$ describes the degradation trend of the battery capacity and the other part $\beta_1 e^{(-\beta_2/\Delta t_k)}$ represents the regenerative capacity during the rest time Δt_k. However, due to the complex operation conditions, the coulomb efficiency η_C doesn't remain constant for each cycle. So, we make η_C as another parameter that needs to be estimated.

With outstanding state tracking capability, the EKF algorithms can be used to identify the parameters in the empirical degradation model. Thus, β_1, β_2 and η_C are the states need to be estimated. Given the empirical model and the capacity degradation data sets, we can conduct the state and measurement updates as described in Eq. (10).

$$\begin{cases} \eta_{C,k} = \eta_{C,k-1} + w_{\eta_c} \\ \beta_{1,k} = \beta_{1,k-1} + w_{\beta_1} \\ \beta_{2,k} = \beta_{2,k-1} + w_{\beta_2} \\ C_{k+1} = \eta_{C,k} C_k + \beta_{1,k} e^{(-\beta_{2,k}/\Delta t_k)} + v_k \end{cases} \tag{10}$$

Here, w_{η_C}, w_{β_1}, w_{β_2} and v_k stand for the noise in the state space model which used for the parameters estimation.

3.2. NSDP-AR Model for Measurements Obtaining

According to the related researches, the ND-AR model has some limitations because the nonlinear degradation is described using an acceleration factor related to the prediction step. This is not reasonable as the parameters of the factor will change a lot with different degraded speed and trajectory, different prediction starting points, sample length of the data, sample interval and density. This makes the ND-AR model not general for other applications.

According to the degradation curve, we can draw a conclusion that the degradation trend has a close relationship with the number of the charging and discharging cycles. The degradation degree increases with the growth of the cycle number. However, the life-cycle length and prediction starting point is different according to individual application. So, we must define or extract a new parameter which can stand for the degradation period with better generalization. Current percentage of the life-cycle length (CPoL) is a scale quantity related to the charging and discharging cycles. Here, CPoL is similar with the definition of the state of health (SOH) of the battery. But the standard definition equations of SOH often use capacity or the power of the battery, so there are certain differences between the two variables. CPoL is the ratio of the current cycles the battery has experienced k and whole cycles during the entire lifetime of the battery L. The specific expression is defined as Eq. (11).

$$CPoL = \frac{k}{L} \tag{11}$$

Eq. (11) shows a better applicability for the individual battery. The CPoL can represent the current internal reaction phase which determines the capacity degradation feature.

However, L is a parameter just needed to be predicted, moreover, the accurate CPoL cannot be obtained in practical applications. Therefore, we need to find an approximated method to estimate it. With GCA analysis, current percentage of the predicted-life-cycle length (CPoP) shown in Eq. (12) indicates high correlation relationship with the CPoL by replacing the L with the predicted-life-cycle length L' obtained using AR model. So, we apply the CPoP as the nonlinear degradation factor, regarding CPoP as an approximated CPoL to implement the proposed approach.

$$CPoP = \frac{k}{L'} \tag{12}$$

The nonlinear degradation factor in this paper has two specific forms. The first one is the same as the related ND-AR model researches have used which can be described as Eq. (5). The other one is obtained from the capacity degradation feature perspective which can be described in an exponential related form (He, Williard, Osterman & Pecht, 2011) (Miao, Xie, Cui, Liang & Pecht, 2012). Two specific forms can be described as Eqs. (13) and (14).

$$K_T = \frac{1}{[1 + a \cdot (kp + b)]} \tag{13}$$

$$K_T = a \cdot e^{b \cdot kp} + c \cdot e^{d \cdot kp} \tag{14}$$

Here kp is the parameter of the CPoL and a, b, c and d are the parameters need to be estimated. With the true degradation data set of certain samples and the corresponding AR model prediction results, we apply EKF algorithm for estimating the parameters in Eqs. (13) and (14). The detailed modeling steps are similar as the ND-AR model (Liu, Luo, Peng, Peng & Pecht, 2012).

In addition, we must determine the parameters of the factor in applications when the true degradation information is unknown. We can conceive that the correlation degree of the history degradation knowledge contributes to the extension of the modified model. Therefore, the weight is determined by the correlation degree and the weighted average results can be described as Eq. (15)

$$m = \sum_{i=1}^{n} m_i \cdot \frac{r_i}{\sum_{j=1}^{n} r_j} \tag{15}$$

where n is the number of batteries we apply for modeling, r_i is the correlation degree between the ith battery sample whose parameter is m_i.

With this improved data-driven method, we can achieve the long-term degradation trend prediction of the capacity.

3.3. Fusion-data-model Prognostic Framework for Lithium-ion Battery RUL Estimation

By applying empirical degradation model as the state transition function with model-based EKF algorithm, and the prediction result with data-driven NSDP-AR model as the measurements update, we can get a novel fusion-data-model prognostic framework shown as Eq. (16).

$$\begin{cases} C_{k+1} = \eta_C \cdot C_k + \beta_1 \cdot e^{(-\beta_2/\Delta t_k)} + w_k \\ y_k = C_k + v_k \end{cases} \tag{16}$$

The whole prognostics framework is shown in Figure 2.

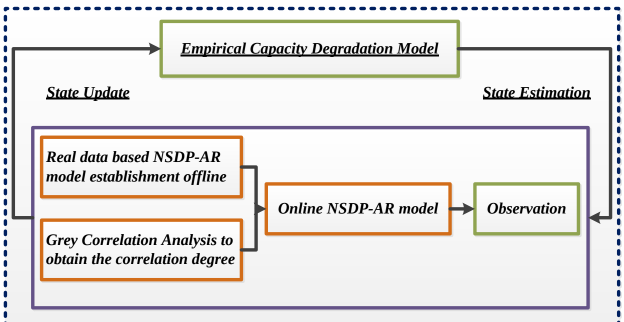

Figure 2. Fusion-data-model Framework for Lithium-ion Battery RUL Estimation

The detailed flow of fusion-data-model is as follows.

Definition: prediction starting point *T*, cycle number *k*, the length of the data for modeling *Length*, real capacity value *Capacity*, predicted capacity value *C*(*k*), failure threshold *U*, long-term trend of the capacity obtained by NSDP-AR model *NSDP-ARpredict*(*k*), and the predicted life-cycle length given by AR model *L'*.

Step 1. Choose individual batteries as the historical samples to modeling the specific NSDP-AR models.

Step 2. Pre-process the data like smoothing, outliers removing and then set parameters such as *T*, *U*, noise covariance matrix and other feature matrix. Capacity data set is divided into the historical part *Capacity*(1:T) and the testing part according to the parameter *T*.

Step 3. Predict with AR model.

Step 3-1. Determine the order of the AR model *p* according to the AIC (Akaike information criterion) principle

$$AIC(p) = N \ln \sigma_p^2 + 2p \tag{17}$$

Here *N* is the number of the data samples, *p* is the model order, σ_p^2 is the prediction variance of *p* order model.

Step 3-2. Use a fusion approach with both Burg method and Yule-Wallker method to calculate the regression coefficient $\phi_i (i = 1, 2, ..., p)$ of the *p* order AR model.

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + a_t \tag{18}$$

Choose the dynamic weight P_{1i} and P_{2i} for the respective results from each method ϕ_{1i} and $\phi_{2i} (i = 1, 2, ..., p)$, then calculate the fusion value as Eq. (19).

$$\phi_i = P_{1i}\phi_{1i} + P_{2i}\phi_{2i} (i = 1, 2, ..., p) \tag{19}$$

Step 3-3. Estimate RUL using AR(*p*) model by performing a multi-step iterative computation to obtain the long-term trend of the capacity *ARpredict*, and the corresponding life-cycle prediction result *L'*.

Step 4. Model NSDP-AR.

Step 4-1. Calculate the real value sequence of the nonlinear degradation factor $K_{T,real}$ using Eq. (20).

$$K_{T,real} = \frac{Capacity}{ARpredict} \tag{20}$$

Step 4-2. Compute the CPoL with Eq. (12) and CPoP with Eq. (11), and ensure the feasibility of the approximated method using GCA algorithm.

Step 4-3. Identify the parameters of the proposed factor with EKF algorithm.

Step 4-4. Model the NSDP-AR by combining the factor and basic AR model for historical individual battery sample.

Step 4-5. Repeat the Step 4-4 for each battery sample.

Step 4-6. Utilize GCA algorithm and weighted average method with Eq. (15) to obtain the NSDP-AR model.

Step 5. Establish the fusion prognostic model with Eq. (16).

Step 6. Predict RUL using the model obtained in Step 5.

There is a key problem need to be explained, that is in this research we only know the data obtained from the system, when we conduct the experiments, the state estimation result will be put into next state estimation step. That is, the prediction is a multi-step iterative prediction, we use the last data point of the collective data set and then all the steps are completed with the help of the previous estimation results.

4. Experiments and Analysis

4.1. Battery Data Set

We utilize two battery data sets from NASA PCoE and CALCE for evaluate the performance of the proposed framework, respectively. The first Li-ion battery data set including battery #5, #6 and #18 are from NASA PCoE. These batteries were tested under certain condition (with the temperature +25℃). The 2 Ah batteries were charged with the charging current 1.5A until the batteries voltage reaches 4.2V, then discharged with the discharging current 2A until the batteries voltage reached 2.5V (Saha, Goebel, Poll & Christophersen, 2009) (Saha & Goebel, 2007). When the battery capacity reaches about 70% of rated capacity, the Li-ion battery is regarded to reaching its end of life (EOL). In the experiments, the threshold is set as 1.38Ah.

The second Li-ion battery data sets are from the CALCE of the University of Maryland containing battery #8, #21 and #33. The cycling of the batteries was implemented with the Arbin BT2000 battery testing system under room temperature. The 1.1 Ah rated capacity of batteries were adopted in the test with the discharging current (0.45A that the discharging speeds is 0.5C) (He, Williard, Osterman & Pecht, 2011) (He, Williard, Osterman & Pecht, 2011) (He, Williard, Osterman & Pecht, 2011). The threshold is set as 0.88 Ah.

4.2. Methods for Comparison

In order to verify the effectiveness of the NSDP-AR model and the fusion prognostic framework, another two algorithms are applied for comparison:

Method 1. Model-based EKF framework for RUL prediction.

Method 2. Fusion-data-model prognostic framework with EKF algorithm and ND-AR model.

These three methods will be utilized to predict the lithium-ion battery RUL.

4.3. Evaluation Criteria

The performance of the algorithm will be evaluated by RUL and capacity prediction errors.

RUL prediction error: We use the RUL prediction result $RUL_{prediction}$ to minus the true RUL value RUL_{real} to describe the RUL prediction error E_{RUL}.

$$E_{RUL} = RUL_{prediction} - RUL_{real} \tag{21}$$

Mean Absolute Error (MAE) for capacity prediction: The MAE is described in Eq. (22) to evaluate the prediction error.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|x(i) - \overline{x}(i)\right| \tag{22}$$

Here, $x(i)$ is the true capacity value and $\overline{x}(i)$ is the prediction value for each cycle, n is the length of data.

Root Mean Square Error (RMSE) for capacity prediction: The RMSE can be described as Eq. (23):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[x(i) - \overline{x}(i)\right]^2} \tag{23}$$

The symbols above have the same meaning as MAE.

4.4. Parameters Identification of the Empirical Degradation Model

With high performance of state tracking for EKF algorithm, we can determine the fit parameters of the empirical degradation model. We have conducted a series of experiments with several data sets to estimate the reasonable model for each battery. In order to describe the tracking result briefly, we choose NASA Battery #18 and CALCE Battery #33 to show the result.

The parameters in the experiments for NASA Battery data are set as follows: the length of the history data set L_1 = 50; the failure threshold U = 1.38Ah; initial state vector [a_0; b_0; c_0] = [1; 10; 10] where a_0 stands for the coulombic efficiency , b_0 and c_0 stand for the unknown parameters β_1 and β_2; the process noise variance matrix Q = [0.0001, 0,0; 0, 0.0001, 0; 0, 0, 0.0001]; observation noise variance R = 0.0001. The weighted average numerical parameters results are: a = 0.9958; b = 10.0040; c = 9.9602.

The parameters in the experiments for CALCE Battery data are set as follows: the length of the history data set L_1 = 280; the failure threshold U = 0.88 Ah; initial state vector [a_0; b_0; c_0] = [1; 10; 10] where a_0, b_0 and c_0 have the same meaning as NASA experiments; the process noise variance Q = [0.0001, 0, 0; 0, 0.0001, 0; 0, 0, 0.0001]; observation noise variance R = 0.0001. The weighted average numerical parameters results are: a = 0.9991; b = 10.0015; c = 9.9847.

4.5. Parameters identification of NSDP-AR Model

According to the modeling flow of the NSDP-AR model, we can obtain the specific nonlinear degradation factor for each battery data set as well as the weighted average result on the basis of the correlation degree of the degradation trend. Both of the factor forms proposed before have been adopted into the experiments respectively for a comparison between different factor forms.

In the experiments for NASA battery data set, Battery #5 and #6 are selected to establish the NSDP-AR models. The obtained weighted parameters will be adopted into the RUL estimation for Battery #18. Similarly, in the experiments for CALCE data set, Battery #8 and #21 are used for modeling and Battery #33 is used for verification. We set the prediction starting point T in the medium-term of each battery sample. Correlation degree between the modeling batteries and the estimated battery are calculated with GCA method where the distinguish coefficient is set as 0.5463. The results are shown in Table 1.

Table 1. Weighted parameters for prediction application

Index	GCA	Form	Parameters			
			a	*b*	*c*	*d*
NASA #5	0.6117	1	-0.0303	0.3938	-	-
NASA #6	0.7612	2	0.9896	0.0028	0.0005	3.9656
CALCE #8	0.6985	1	-0.1326	-0.5346	-	-
CALCE #21	0.6776	2	1.0357	-0.0268	0.0011	3.3169

4.6. RUL Estimation with the Fusion-data-model Prognostic Framework

With the parameters obtained in Section 4.4 and 4.5, we conducted RUL estimation experiments using different methods. With the fusion prognostics, the RUL prediction results of NASA Battery #18 are shown in Figures 3 and 4. The result in Figure 3 is based on the nonlinear degradation factor form as Eq. (13), while Figure 4 shows the result based on the factor as Eq. (14).

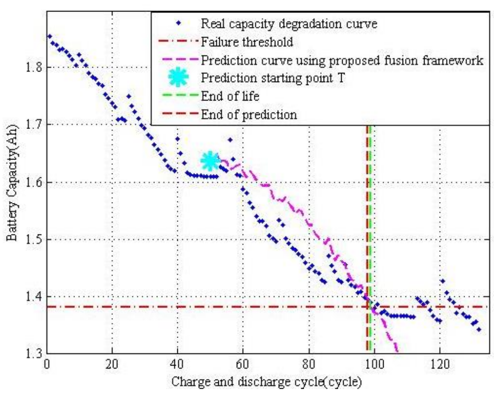

Figure 3. RUL prediction based on proposed fusion framework for NASA Battery #18 (Form 1)

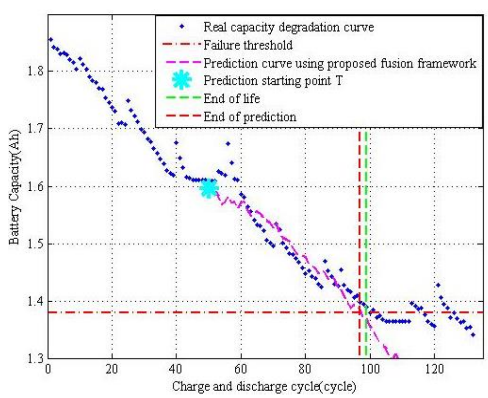

Figure 4. RUL prediction based on proposed fusion framework for NASA Battery #18 (Form 2)

Similarly, Figures 5 and 6 show the experimental results for CALCE battery data sets. Here, two different factor forms are also adopted in the fusion prognostic framework, which are similar as the experiments above.

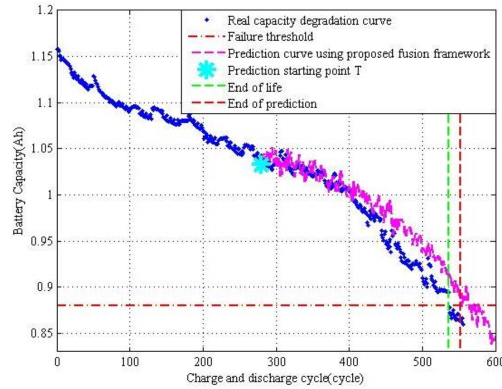

Figure 5. RUL prediction based on proposed fusion framework for CALCE Battery #33 (Form 1)

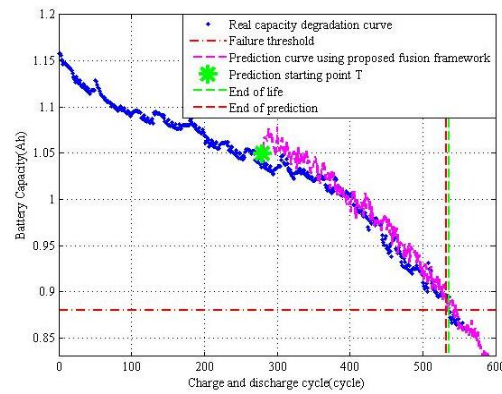

Figure 6. RUL prediction based on proposed fusion framework for CALCE Battery #33 (Form 2)

From the RUL prediction results for two types of battery data sets, we can find that the degradation curves as well as the estimated EOL points are very close to the actual values. It is indicated that the proposed method is efficient and accurate.

4.7. Comparison and Analysis

To compare and evaluate different RUL prediction methods, we also conducted more experiments. The quantitative results are shown in Table 2 and Table 3 for NASA and CALCE batteries, respectively.

By analyzing the experimental results above, it can be seen that the proposed method can achieve better performance than the other methods with the MAE and RMSE shown in Tables 2 and 3. The ND-AR model based on prediction step k may sometimes make a deterioration of the prediction results, because this method does not take into account the diversity of degradation speed. In contrast, the NSDP-AR model can overcome this problem and lead to a better RUL estimated result. It is showed that the NSDP-AR model based fusion framework achieve an improvement on prediction accuracy.

Table 2. Comparison of RUL estimation with different prognostics for NASA Battery #18

Methods	E_{RUL}	MAE	RMSE
EKF	9	0.0804	0.1112
EKF & ND-AR (1)	1	0.0862	0.1335
EKF & ND-AR (2)	5	0.0458	0.0636
EKF & NSDP-AR (1)	1	0.0938	0.1338
EKF & NSDP-AR (2)	2	0.0726	0.1083

Table 3. Comparison of RUL estimation with different prognostics for CALCE Battery #33

Methods	E_{RUL}	MAE	RMSE
EKF	0	0.0206	0.0239
EKF & ND-AR (1)	48	0.0287	0.0357
EKF & ND-AR (2)	×	116.5041	257.2003
EKF & NSDP-AR (1)	16	0.0173	0.0213
EKF & NSDP-AR (2)	4	0.0133	0.0162

Note that, the RUL estimated error for CALCE Battery #33 obtained with EKF algorithm equals to 0 in Table 3. However, the estimated degradation trend is obviously diverse from the real curve, and the satisfied RUL prediction result is a coincidence for certain battery samples.

At the same time, the real-time performance of EKF algorithm is superior to the other statistical filtering such as Particle Filter. The comparison between the operating speed of PF and EKF is shown in Table 4. We can find that the execution time of fusion EKF algorithm is shorter than the fusion PF method but the accuracy is close with each other.

Table 4. Comparison of RUL estimation with different prognostics for CALCE Battery #33

Methods	E_{RUL}(cycle)	Time(s)
Fusion EKF	7	2.27
Fusion PF	9	11.45

Moreover, the NSDP-AR model realizes satisfied prediction for the nonlinear degradation trend, and more reasonable than ND-AR model, which stands that the proposed fusion framework shows a good application prospect.

5. Conclusions

This paper explores an improved fusion-data-model prognostic framework with EKF algorithm and NSDP-AR model for battery RUL estimation. The main contributions of this research can be concluded as follows. (1) A data-model fusion prognostic framework with low computation complexity and better real-time capability is proposed. (2) Improvements are obtained to modify the ND-AR model including a scale parameter based nonlinear degradation analysis, and an approximate method to obtain the CPoL data sets, and an extension method using GCA method, as a result, an improved NSDP-AR model is achieved. (3) A combination of model-based EKF algorithm and improved data-driven NSDP-AR model which weakens the dependence on empirical degradation model and improves the nonlinear predicting accuracy.

6. Future Work

In future, more efforts should be focused on obtaining the RUL estimation result when the capacity cannot be observed on-line, which is the actual condition in practical applications. Up to now, we have conducted some studies with a novel parameter called time intervals to equal discharging voltage difference (TIEDVD) to realize indirect RUL prognostics. Meanwhile, in this paper we only conduct the study on estimating the RUL when the battery is fully discharged and discharged, which is not true in practical applications. On the other hand, the empirical degradation model used in this framework only takes the capacity degradation principle and the regeneration phenomenon of the capacity during test time into consideration. More factors such as depth of discharge (DOD) and internal temperature of the lithium-ion battery should be considered while modeling. Especially, the indirect health indicator (HI) with on-line monitoring parameters and the fusion HI combines with more correlation analysis methods should be applied to realize more flexible and applicable RUL estimation.

Acknowledgement

This work is partly supported by National Natural Science Foundation of China under Grant No. 61301205, Twelfth Government Advanced Research Fund under Grant No. 51317040302, Research Fund for the Doctoral Program of Higher Education of China under Grant No. 20112302120027, Fundamental Research Funds for the Central Universities under Grant No. HIT.NSRIF.2014017, and China Scholarship Council. The author would like to thank Dr. Zhimin Xi at University of Michigan – Dearborn for his revision and suggestion to this manuscript.

References

He, W., Williard, N., Osterman, M., & Pecht, M. (2011). Prognostics of Lithium-ion Batteries using Extended Kalman Filtering. *IMAPS Advanced Technology Workshop on High Reliability Microelectronics for*

Military Applications. Linthicum Heights, MD, USA, 1- 4.

He, W., Williard, N., Osterman, M., & Pecht, M. (2011). Remaining useful performance analysis of batteries. *2011 IEEE Conference on Prognostics and Health Management* (*PHM*) *Conference*, Denver, CO, USA, June 20-23.

He, W., Williard, N., Osterman, M., & Pecht, M. (2011). Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method. *Journal of Power Sources*, vol. 196, pp. 10314-10321.

Liu, D., Luo, Y., Peng, Y., Peng, X., & Pecht, M. (2012). Lithium-ion Battery Remaining Useful Life Estimation Based on Nonlinear AR Model Combined with Degradation Feature. *Annual Conference of the Prognostics and Health Management Society 2012*, Minneapolis, Minnesota, USA, September, 24-27.

Liu, J., Saxena, A., Goebel, K., Saha, B., & Wang, W. (2010). An Adaptive Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-ion Batteries. *Annual Conference of the Prognostics and Health Management Society 2010*. Portland, Oregon, USA, October 10-16.

Liu, J., Wang, W., Ma, F., Yang, Y. B., & Yang, C. S. (2012). A data-model-fusion prognostic framework for dynamic system state forecasting. *Engineering Applications of Artificial Intelligence*, vol. 25, pp. 814-823.

Miao, Q., Xie, L., Cui, H., Liang, W., & Pecht, M. (2012). Remaining useful life prediction of lithium-ion battery with unscented particle filter technique. *Microelectronics Reliability*: vol. 53, pp. 805-810.

Saha, B., & Goebel, K. (2007). "Battery Data Set", NASA Ames Prognostics Data Repository, [http://ti.arc.nasa.gov/project/prognostic-data-repository], NASA Ames, Moffett Field, CA.

Saha, B., Goebel, K., Poll, S., & Christophersen, J. (2009), Prognostics Methods for Battery Health Monitoring Using a Bayesian Framework. *IEEE Transactions on Instrumentation and Measurement*, vol. 58, pp. 291-297.

Saha, B., & Goebel, K. (2009). Modeling li-ion battery capacity depletion in a particle filtering framework. *Annual Conference of the Prognostics and Health Management Society 2009*, San Diego, CA, USA, September 27 – October 1.

Shen, M., Xue, X., & Zhang, X. (2003). Determination of Discrimination Coefficient in Grey Incidence Analysis. *Journal of Air Force Engineering University* (*Natural Science Edition*), vol. 4, pp. 68-70.

Wei, W. W. S. (1994). *Time series analysis*. Redwood City, California: Addison-Wesley.

Welch, G., & Bishop, G. (1995). *An introduction to the Kalman filter*.

Zhang, J., & Lee, J. (2011). A review on prognostics and health monitoring of Li-ion battery. *Journal of Power Sources*, vol. 196, pp. 6007-6014.

BIOGRAPHIES

Datong Liu received the B.Sc. and M.Sc. degrees in Department of Automatic Test and Control from Harbin Institute of Technology (HIT), Harbin, China in 2003 and 2005, respectively. During 2001 to 2003, he also minored in the Computer Science and Technology in HIT. He received the Ph.D. degree in major of measurement and instrumentation from HIT in 2010. He is now an assistant professor in Department of Automatic Test and Control, HIT. His research interests include automatic test and intelligent information processing, time series analysis, Data-driven PHM, Machine Learning, Data Mining, etc. He is currently an IEEE member, ACM member, PHM society member, China Computer Federation member. He has published more than 30 journal and conference papers, and holds 15 invention patents and more than 30 invention patents pending in China. He is now in charge of 8 projects related to PHM that supported by National Natural Science Foundation of China, Research Fund for the Doctoral Program of Higher Education of China, Twelfth government advanced research fund in China, Fundamental Research Funds for the Central Universities, etc.

Limeng Guo received her B.Sc. degree in Department of Automatic Test and Control from Harbin Institute of Technology (HIT), Harbin, China in 2013. Now she is a postgraduate in Major of Instrumentation Science and Technology of HIT. Her research interests include data-driven PHM, fusion prognostic approach, Battery Management System, and so on. She wins the best paper award of ICEMI 2013, and best graduated thesis award of HIT in 2013. She is now involved in 5 projects focusing on system health management of complex system, and also holds 10 invention patents pending in China.